Personality and Individual Differences 229 (2024) 112751

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Developmental trends in intelligence revisited with novel kinships: Monozygotic twins reared apart *v.* same-age unrelated siblings reared together

Nancy L. Segal *, Elizabeth Pratt-Thompson

Psychology Department, California State University, Fullerton, CA 92831, USA

ARTICLE INFO

Keywords:
Twins
Adoptees
Siblings
IQ
Intelligence
Genetics
Environment

ABSTRACT

Prior research has reported convergence of monozygotic (MZ) cotwins' intelligence scores over time, with divergence observed between dizygotic (DZ) cotwins. These patterns have been variously explained with reference to the increased influence of genetic factors (MZ and DZ twins) and nonshared environmental factors (DZ twins), and the reduced effects of shared environmental events (DZ twins). Studies of unrelated siblings have found modest within-pair resemblance in intelligence, with increasing divergence over time. This subject is revisited in a study using three novel kinships: young reared-apart MZ twins from China (CTA-MZ), adult reared-apart twins from Denmark (D-MZA) and findings from a previous study of same-age unrelated siblings (virtual twins or VTs). Despite modest sample sizes, the anticipated trends described above were observed, replicating extant findings. Intraclass correlations for overall IQ score at Time 1 and Time 2, respectively, were r_is = 0.51, 0.81 (CTA-MZs) and r_is = 0.64, 0.74 (D-MZAs). The hypothesis that VTs would show score divergence at Time 2 was confirmed. Increased genetic influence (CTA-MZ, D-MZA and VT), reduced impact of shared environments (VT), and increased effects of nonshared environments (VT) appear to best explain the findings. These results also inform parent and educator expectations regarding twins' and siblings' academic performance.

1. Introduction

1.1. Intellectual landscape: nature and nurture

Factors affecting the development of general intelligence have engaged the interest of behavioral science professionals, educators, and members of the public for years. A meta-analysis of twin studies conducted between 1967 and 1985 by McCartney, Harris, and Bernieri (1990) failed to find an age effect for the IQ resemblance of MZ and DZ twins raised together. However, an important limitation of this analysis was that the study was cross-sectional in nature; the twins were not tested at more than one time point. Indeed, the authors noted the contrast in findings between their study and the landmark work of Wilson (1978), more recently labelled the *Wilson Effect* (Bouchard, 2013). Using IQ data gathered on reared-together twins tested periodically between three months and fifteen years of age, Wilson (1978) found that MZ twins showed increasing IQ convergence, in contrast with DZ twins who showed increasing IQ divergence. The McCartney et al. (1990) study proposed that the unique social dynamics between MZ twins raised together may partly explain their findings. That is, they reasoned that efforts toward social differentiation between MZ co-twins may have contributed to their somewhat different intellectual outcomes.

Bouchard (2013) reinforced Wilson's finding in his comprehensive review of twin and adoption studies of general intelligence, conducted across different countries with different populations and protocols. Even the somewhat lowered heritability of 0.62 evidenced by 80-year-old reared-together Swedish twins (McClearn et al., 1997), relative to a heritability of 0.77 based on adult reared-apart twins from the Minnesota Study of Twins Reared Apart, is persuasive of genetic effects (Johnson et al., 2007). Consistent with these findings is the reduced and diminishing resemblance of adopted siblings assessed over time (Loehlin, Horn & Willerman, 1989; Scarr, Weinberg, & Waldman, 1993; Segal, McGuire, Havlena, Gill, & Hershberger, 2007). Collectively, these studies variously highlight the increasing effects of emerging genetic factors, the importance of nonshared environments, and the reduced effects of shared experiences on shaping intellectual development.

* Corresponding author at: CSU Fullerton, Psychology Department, 800 N. State College Blvd., H830M, Fullerton, CA 92831, USA.
E-mail address: nsegal@fullerton.edu (N.L. Segal).
URL: https://www.drnancysegaltwins.org (N.L. Segal).

https://doi.org/10.1016/j.paid.2024.112751
Received 25 March 2024; Accepted 1 June 2024
Available online 15 June 2024

Surprisingly, an analysis of reared-apart twins IQ-tested at more than one time point has not appeared in the extant literature, with the exception of Juel-Nielsen's classic (1980) study. Juel-Nielsen did not detect a meaningful age effect on the twins' mean IQ score upon dividing the pairs into those above and those below fifty-five years of age. This is not surprising, given that all participants were adults at both test sessions, such that their general intelligence level would have been stable. It is curious that Juel-Nielsen did not compute the Time 2 IQ intraclass correlation for comparison with the correlation at Time 1, an analysis that was conducted in the present study. It is also unfortunate that the Raven Progressive Matrices test was administered to the Danish twins at Time 1, but not at Time 2, as such a comparison would have relevance to the hypotheses currently under investigation.

Perhaps the most important gap in the relevant literature is the absence of a reared-apart twin analysis of developmental changes in IQ using children. Such an undertaking would provide findings more comparable to those from the few adoptive studies that have followed subjects from childhood through adolescence. Thus, the present study is the first to do so, using a unique sample of mostly young Chinese twins who grew up apart. The rarity and expense of gathering longitudinal data have been emphasized, as has the value of obtaining data at multiple time points. As Lykken (2007) noted, psychological measures vary around their stable set-points due to environmental fluctuations, such that data collected on even two occasions offers more accurate estimates of heritability. Findings from such a study, contrasting CTA-MZA pairs (Chinese MZ twins reared apart) and D-MZA pairs (Danish twins reared apart) with previously reported findings from an exceptional subset of unrelated sibling pairs, i.e., virtual twins (VTs: same-age unrelated siblings reared together who replicate twinship, but without a genetic link), are presented.

1.2. Hypotheses

(a) CTA-MZ and D-MZA Twin Pairs: IQ similarity will increase and within-pair IQ score differences will decrease from Time 1 to Time 2.

(b) The interval (time) between Time 1 and Time 2 will be uncorrelated with the IQ within-pair difference scores of the CTA-MZ and D-MZA twin pairs.

(c) Age at separation and contact time between Time 1 and Time 2 will be uncorrelated with the IQ within-pair difference scores of the reared-apart twins.

(d) IQ heritability derived from the reared-apart twins, and environmental effects based on the virtual twins will be higher using two IQ measures, compared to using one.

2. Materials and method

2.1. Participants

The members of three unique participant groups comprised the samples for the present study. Descriptions of these samples are provided below.

2.2. CTA-MZ twins

The first sample in the present study included young MZ reared-apart twins, mostly from China, who were separated as an indirect consequence of that nation's One-Child Policy (CTA). This policy, in effect from 1980 to 2016, limited urban families to one child and rural families to two children (Buckley, 2015). Many pregnant women underwent forced sterilizations and abortions, while others abandoned their children on the steps of orphanages and police stations (Evans, 2008). Given the preference for male children in Chinese culture, the majority of abandoned children were female, and among them were twins; the complete project sample included only one male twin from an opposite-sex pair. Twins in two pairs from Vietnam and twins in one pair from Taiwan were reared apart due to adverse family circumstances, e.g., single parent stigma and/or insufficient financial resources.

Both monozygotic (MZ) and dizygotic (DZ) twins have been enrolled in the *Fullerton Study of Chinese Twins Adopted Apart and Together* since its inception in 2006. Reared-apart twins in the present study were variously identified through the media (63.5 %), referrals (13.6 %), self-referrals (18.2 %), and other means (4.7 %). Parents of twins received an informed consent letter and a packet of demographic questionnaires and behavioral inventories to complete for their children. Several twins older than eighteen years of age received similar materials; parents' responses were used for respondent consistency, with the exception of contact time as reported by older pairs at Time 2.

The first sample includes MZ reared-apart twins only (CTA-MZ). These pairs offer simple, elegant, and convenient control of genetic and environmental effects when contrasted with other genetically and environmentally informative kinship pairs, as explained below. The CTA-MZ sample ($N = 15$ pairs) had a mean age of $\overline{x} = 10.69$ years ($SD = 7.20$) at Time 1, and a mean age of $\overline{x} = 13.93$ ($SD = 6.91$) at Time 2. (One pair with incomplete data was omitted from the Time 2 analyses). Other features important to the present study, e.g., age at adoption and test-retest interval, are summarized in Table 1a ($N = 30$ individuals). Additional information about the origins and progress of the study is available in Segal, Niculae, Becker, and Shih (2021) and references therein. The CTA-DZ twin sample was too modest in size to provide meaningful IQ findings.

2.3. D-MZA twins

A detailed volume describing the origins, methods, participants, results, and conclusions of a Danish study of adult reared-apart twins was the source of IQ data used in the present analysis (Juel-Nielsen, 1980). The twelve twin pairs comprising the sample are unique in that they represent the entire population of separated sets in Denmark, identified between 1954 and 1957. The first set came to the investigator's attention via his association with the psychiatric unit of the State Hospital in Riskov. The second set was referred to him by a journalist covering the twins' story. The remaining ten sets were identified through the twins registry of the Institute of Human Genetics at the University of Copenhagen. Their mean age at first participation was $\overline{x} = 51.42$ years, $SD = 16.70$.

The study was a comprehensive medical and psychological investigation. In addition to general intelligence testing, assessments were made of personality, general health, physical traits, and early environmental differences. The quantitative findings are supplemented with detailed biographical sketches of each separated pair. One or both cotwins in three pairs did not complete the second wave of data collection, leaving nine complete pairs ($N = 18$ individuals) for examining the IQ-related hypotheses listed above. Age, age at separation and test-retest interval are displayed in Table 1b. Some data displayed in the tables and used in the analyses differ slightly from Juel-Nielsen's, given that

Table 1a
Age, age difference at testing, and age at adoption of CTA-MZ twins at Time 1 and Time 2 [$N = 14$–15 pairs]; ($N = 28$–30 individuals).

Measure	Time	Mean	SD	Range
Age at testing in years				
($N = 30$)	1	10.69	7.20	3.19–24.98
($N = 28$)	2	13.93	6.91	5.42–29.95
Age difference at testing in days				
[$N = 15$]	1	12.73	10.86	0.00–41.00
[$N = 14$]	2	0.09	0.10	0.01–0.40
Age at adoption[1] in months ($N = 30$)		19.60	38.30	4.20–218.20
Test-retest interval in years [$N = 14$]		4.18	2.04	2.02–9.12

(N) = 30 individuals [N] = 14 pairs or families.

Table 1b
Age, age at separation, and test-retest interval of adult Danish twins (D-MZA) at Time 1 and Time 2.

Measure	Time	Mean	SD	Range
Age at testing in years[a]	1	51.42	16.70	22.00–77.00
[N = 12]	2			
Age at separation in Months [N = 12]		16.42	21.53	0.03–66.00
Test-retest interval[b] in months (N = 18)		11.17	(6.57)	6.00–31.00

(N) = 18 individuals. [N] = 9 pairs or families.
[a] Age at second testing not given.
[b] The mean test-retest interval was based on 9 complete twin pairs.

only complete pairs were included in the present study.

2.4. VTs

Detailed information about the VTs is provided in the original paper (Segal et al., 2007), so only selected features of this kinship are summarized here. VTs are same-age unrelated siblings reared together soon after birth that can be organized into one of two main types: adopted-adopted or adopted-biological. Defining criteria for virtual twinship are age difference of nine months or less, entry into the home by age one year, enrollment in the same school grade, and absence of adverse birth events that would affect cognitive performance (Segal, 1997). Thus, VTs afford sensitive comparisons with MZ twins, regardless of twin' rearing status. Specifically, VTs' matched age and closely timed entry into the home distinguish them from ordinary adoptive sibling pairs—as indicated, VTs replay twinship, but without genes shared by descent (Segal, Tan, & Graham, 2015). In the comparison presented here the VTs provide a direct estimate of shared environmental effects, whereas the CTA-MZs and D-MZAs provide a direct estimate of genetic effects. Age, age differences in testing, age at adoption, and test-retest interval for the VTs twin pairs are displayed in Table 1c.

The VTs (N = 43 pairs) were a subset of the sample of 170 pairs currently enrolled in the Fullerton Virtual Twin Study. The mean age of the VTs (N = 84 individuals; 2 individuals were in more than one pair) was $\overline{x}$= 5.11 years (SD = 1.10) at Time 1 and $\overline{x}$ = 10.77 (SD = 1.61) at Time 2. The sample included 30 adopted-adopted pairs (70 %) and 13 adopted-biological pairs (30 %). Additional information about the Fullerton Virtual Twin Study is available in Segal (2006), Segal et al. (2007), Segal and Niculae (2019), and Segal et al. (2021).

2.5. Procedures and protocols

2.5.1. Zygosity diagnosis (twins)

The monozygosity of 14 of the 15 CTA-MZ twin pairs was established by DNA analysis derived from buccal smears (saliva samples). These assignments aligned with results from a standard, widely used physical resemblance questionnaire, devised by Nichols and Bilbro (1966) and completed by the twins' mother. Cross-validation in the composition of the form showed that approximately 95 % of the development sample could be diagnosed with greater than 90 % accuracy. The single pair in the present sample that did not undergo DNA testing was classified as MZ, based on questionnaire scores. The zygosity of the D-MZA twins was assigned by resemblance in the highly heritable physical traits of blood/serum groups, eye color, hair color and dermatoglyphic features.

Table 1c
Age, age difference at testing, age at adoption and test-retest interval of virtual twins at Time 1 and Time 2 [N = 43 pairs]; (N = 84 individuals).

Measure	Time	Mean	SD	Range
Age at testing in years	1	5.11	1.10	4.00–8.70
(N = 84)	2	10.77	1.61	7.18–13.67
Age difference at testing in months	1	3.05	2.62	0.00–8.93
[N = 43]	2	3.02	2.64	0.00–8.93
Age at adoption[a] in months (N = 84)		1.09	2.21	0.00–9.87
Test-retest interval in years (N = 84)		5.65	1.68	1.70–8.96

(N) = 84 individuals; one adoptee paired with triplets was counted only once in. Individual analyses [N] = 43 pairs or families.
[a] Age at entry into the family for adoptive children; age at hospital release for biological children.

2.5.2. General intelligence testing

2.5.2.1. *CTA-MZ.* The age-appropriate Wechsler IQ test was administered locally to the members of CTA-MZ pairs by different trained testers who were blind to the hypotheses and questions under investigation. Upon receipt, scoring was reviewed by the PI, allowing questions to be resolved prior to data processing. Efforts were made to minimize cotwins' interval between the two test sessions.

The use of separate testers for each pair member is a prudent practice in twin and sibling research, a field in which genetic findings have been challenged from time to time (Segal, 2012). However, such precautions made in the interest of avoiding biased administration and scoring appear unnecessary if the rules specified in the test manuals are closely followed (Segal & Russell, 1991).

2.5.2.2. *D-MZA.* The Danish twins completed the Wechsler-Bellevue test of intelligence-Form I (Wechsler, 1939). Cotwins were tested independently, but at the same time to avoid exchange of information. It appears that the same investigator tested both twins initially, but without knowledge of their biographical histories. Retesting was performed by this same examiner for six pairs and for one-cotwin in another pair. A second examiner administered the test to both co-twins in two pairs, yielding nine complete pairs who completed the test on two occasions.

2.5.2.3. *VTs.* The members of forty-one of the forty-three VT sets (eighty-two individuals) completed a second intelligence test during their participation in TAPS (Twin, Adoptees, Peers, and Siblings), a project that examined intelligence, friendship, adjustment and parenting of twins, full siblings, adoptees, and friends (McGuire & Segal, 2013). Cotwins in two pairs were tested twice as part of the original Fullerton study (Segal et al., 2007). Like CTA-MZs, the VTs completed the age-appropriate Wechsler IQ test. Recall that VTs resided in the same home. Therefore, with only a few exceptions, siblings were routinely tested by different examiners on the same day at both Time 1 and Time 2 to avoid biased administration and/or scoring; but see above. Time 2 scores for VTs were calculated from two verbal subtests (Information and Vocabulary) and two nonverbal subtests (Block Design and Picture Arrangement) from the WISC-III. The composite score derived from this process correlates higher than r = 0.90 with the IQ score derived from the complete protocol (Sattler, 1989).

2.6. Analytical procedures

Data analysis was organized into two parts: Part I. Descriptive features of the Time 1 and Time 2 IQ scores (e.g., means, standard deviations, range, and test-retest correlations) were obtained and compared. Intraclass correlations for IQ scores and the mean within-pair difference score were calculated and examined with reference to hypotheses a-b. Correlations between age at separation and contact time were examined with reference to hypothesis c. Part II: The cross-twin – cross-time correlation (R_{ct}) and retest correlation (R_{wt}) were used to recalculate IQ heritability (reared-apart twins) and environmentality (virtual twins). The R_{ct}/R_{wt} ratio estimates the true or disattentuated resemblance of the cotwins'/cosibs' trait set-points (Lykken, 2007).

Table 2a
CTA-MZ Twins' IQ scores at time 1 and time 2.

Time	Mean[a]	(SD)	Range	r (T1,T2)
Full scale IQ score[b]				
1	109.96	(14.31)	79–136	0.83***
2	109.46	(16.82)	71–136	
T2-T1[c]	−0.50	(9.45)	−23–17	
Interval (years)	4.20	(2.03)	2.00–9.00	

[a] Individual Data ($N = 28$).
[b] Time 2 v. Time 1 non-significant.
[c] Signed Mean Difference.
*** $p < .001$.

3. Results

3.1. Part I

3.1.1. Intelligence test scores: CTA-MZ

The mean IQ and within-pair difference scores of the CTA-MZs are displayed in Table 2a. The CTA-MZs' mean Full Scale IQ scores were nearly identical at Time 1 and Time 2. The correlation between tests taken at Time 1 and Time 2 was statistically significant ($r = 0.83$, $p < .001$). The IQ test-retest correlation exceeds values reported in the psychological literature for diverse groups of biological children ($rs = 0.57$ to 0.63), tested at ages 7–8 years and again at 17–18 years (Scarr et al., 1993). A more recent survey reported correlations of 0.47 -. 78 for tests (WPPISI and WISC) taken between young childhood, middle childhood, and adolescence (Deary, Whalley, Lemmon, Crawford, & Starr, 2000). The mean interval (time) between the two test administrations was $\bar{x} = 4.20$ years ($SD = 2.03$).

3.1.2. Intelligence test scores: D-MZA

The mean IQ and within-pair difference scores of the D-MZs are displayed in Table 2b. The D-MZAs' mean Full Scale IQ scores were several points higher at Time 2, a difference that was statistically significant. The test-retest correlation between tests taken at Time 1 and Time 2 also reached statistical significance ($r = 0.88$, $p < .001$); this correlation aligns well with the reliabilities ($rs = 0.90$–0.91) reported in four studies of typical individuals (Derner, Aborn, & Canter, 1950). The mean interval (time) between the two test administrations was $\bar{x} = 11.17$ months ($SD = 6.57$) and was unrelated to the within-pair difference at Time 2. Age at separation was also uncorrelated with the twins' IQ scores and within-pair difference scores at both Time 1 and Time 2.

3.1.3. Intelligence test scores: VT

The VTs' Full Scale IQ score was significantly higher at Time 2 than at Time 1, as reported previously (Segal et al., 2007). The test-retest correlation ($r = 0.50$, $p < .01$) was generally consistent with values reported for diverse child groups, cited above. The mean interval (time) between test sessions was $\bar{x} = 5.65$ years ($SD = 1.68$). These data are displayed in Table 2c.

Table 2b
D-MZA Twins' IQ scores at time 1 and time 2.

Time	Mean[a]	(SD)	Range	r (T1,T2)
Full Scale IQ Score[b]				
1	105.44	(10.31)	91–125	0.88***
2	108.61	(10.08)	96–135	
T2-T1[c]	3.17	(4.93)	−6–14	
Interval (months)	11.17	(6.57)	6–31	

$t(17) = -2.73$, $p < .01$.
[a] Individual Data ($N = 18$).
[b] Time 2 v. Time 1 non-significant.
[c] Signed mean difference.
*** $p < .001$.

Table 2c
Virtual twins' IQ scores at time 1 and time 2.

Time	Mean[a,2]	(SD)	Range	r (T1,T2)
Full scale IQ Score[2]				
1	105.86	(11.41)	83–135	
2	108.89	(13.25)	81–145	
T2-T1[b]	3.03	(12.38)	−28-31	0.50**
Interval (years)	5.65	(1.58)	1.70–8.96	

[a] Individual data ($N = 84$).
[2] $t(83) = -2.25$, $p < .05$.
[b] Signed mean difference.
** $p < .01$.

Table 3a
Chinese reared-apart twins: IQ intraclass correlations, within-pair differences, and 95 % confidence intervals at Time 1 and Time 2.

Time 1	r_i	95 % CI	Within-pair diff[1]	(SD)	Range
Full IQ					
1	0.51*	(0.03, 0.80)	11.93	(7.49)	1–30
2	0.81***	(0.52, 0.93)	7.93	(7.14)	0–22

$N = 15$ pairs, Time 1; $N = 14$ pairs, Time 2.
[1] $t(13) = 2.12$, $p < .05$.
* $p < .05$.
*** $p < .001$.

3.1.4. IQ similarities and differences: hypothesis testing

The IQ intraclass correlation for the CTA-MZ twins showed the anticipated increase in from Time 1 to Time 2 ($r_i = 0.51$ to 0.81), although the difference was not statistically significant. The anticipated reduction in the within-pair difference scores was also observed. These values for CTA-MZs at Time 1 and Time 2, respectively, were $\bar{x}= 11.93$ ($SD = 7.49$) and $\bar{x} = 7.93$ ($SD = 7.14$). These data are displayed in Table 3a. The mean test-retest interval of $\bar{x} = 4.20$ years, $SD = 2.03$ correlated significantly with the absolute within-pair difference ($rs = -0.45$, $p = .018$), but in a counterintuitive direction—larger intervals were associated with smaller within-pair IQ differences. Age at separation and age at adoption were uncorrelated with the IQ within-pair difference at both Time 1 and Time 2. Contact time subsequent to the first test session was unrelated to the IQ within-pair difference score at Time 2.

In addition to the hypotheses assessed above, it was observed that age at adoption correlated significantly with twins' IQ score at both Time 1 ($r = -0.48$, $p < .01$) and Time 2 ($r = -0.38$, $p < .05$). The direction of

Table 3b
Danish reared-apart twins: IQ intraclass correlations, within-pair differences, and 95 % confidence intervals at Time 1 and Time 2.

Time	r_i	95 % CI	Within-pair diff[a]	(SD)	Range
Full IQ					
1	0.64**	(0.15, 0.88)	7.25	(4.03)	1–14
2	0.74**	(0.23, 0.93)	6.56	(3.84)	3–15

$N = 12$ pairs Time 1; $N = 9$ pairs Time 2.
** $p < .01$.
[a] $t(8) = ns$.

Table 3c
Virtual twins: IQ intraclass correlations, within-pair differences, and 95 % confidence intervals at time 1 and time 2.

Time	r_i	95 % CI	Within-pair diff[1]	(SD)	Range
Full IQ					
1	0.30*	(0.01, 0.55)	10.74	(8.31)	0–40
2	0.11	(−0.19, 0.39)	14.12	(10.39)	0–50

($N = 43$ pairs).
[1] $t(42) = -2.03$, $p < .05$.
* $p < .05$.

the correlations indicates that earlier adoption was associated with a higher IQ.

An increase in the magnitude of the intraclass correlation for the D-MZA pairs was noted, but was not statistically significant, nor was the relationship between the test-retest interval and the within-pair difference at the second testing. The within-pair IQ difference showed a slight decrease at Time 2, as shown in Table 3b; unfortunately, this finding could not be evaluated with reference to contact time between test sessions, due to the absence of this information in Juel-Nielsen's original study. However, contact time during "upbringing" was rated as "none," "slight," and "moderate." These ratings were unrelated to IQ scores and within-pair difference scores at both Time 1 and Time 2.

As expected, the test interval was also unrelated to the within-pair IQ difference. Age at separation was unrelated to the within-pair IQ difference at Time 1 and Time 2. However, in contrast with the CTA-MZ twins, a significant increase in IQ score from Time 1 to Time 2 was noted.

3.1.5. IQ similarity and differences: hypothesis testing (VT)

As reported earlier (Segal et al., 2007), the intraclass correlation for IQ showed the anticipated decrease from r_i = -0.30 to 0.11. The anticipated increase in the within-pair difference scores was also observed, a difference that was statistically significant ($p < .05$). These values at Time 1 and Time 2, respectively, were $\bar{x} = 10.74$. (*SD* 8.31) and $\bar{x} = 14.12$ ($SD = 10.39$). The mean test-retest interval showed a negligible correlation with the within-pair IQ difference score. These data are displayed in Table 3c.

3.2. Part II

3.2.1. Heritability and environmentality recalculated

The CTA-MZ correlation of the Time 1 and Time 2 scores (*R* of Means = 0.69) exceeded the Time 1 correlation, but not the Time 2 correlation. In contrast, the comparable D-MZA correlation exceeded the Time 1 and Time 2 correlations, the latter albeit slightly (*R* of Means = 0.76). As expected, the estimated heritability for both reared-apart samples increased when the R_{ct}/R_{wt} was used, compared with the Mean of *Rs* (0.65 *v.* 0.70 and 0.68 v. 0.75, respectively), consistent with Lykken's (2007) discussion. In contrast, the estimated environmentality for the VTs decreased, as expected (R_{ct}/R_{wt} = 0.12), indicating that shared environments contribute less to the IQ similarity of adoptive siblings over time. These findings are summarized in Table 4.

4. Discussion

4.1. Part I findings

The first longitudinal prospective study of young reared-apart MZ twins (CTA-MZ) showed convergence in IQ similarity (increased intraclass correlation and reduced within-pair difference) over time, as anticipated. However, these results require cautious interpretation given the small sample size. The lack of statistical significance in the Time 1 and Time 2 correlations most likely reflects that limitation. Nevertheless, the uniqueness of these data (longitudinal in nature) and the participant group (young reared-apart twins) lends the findings greater meaning than they might have otherwise. Furthermore, reared-apart twins offer significantly greater power than reared-together twins (Lykken, Geisser, & Tellegen, 1981). Despite their different homes, educational experiences, and (in some cases) residences in different countries, the twins appear to have interacted with their environments in ways that aligned with their genetic propensities. This supports the notion that environments do not act randomly in fashioning developmental outcomes—rather, individuals behave selectively and actively with respect to the people, places and events that engage and challenge them.

The foregoing explanation illustrates the simple, but profound concept of *nature* via *nurture* (Bouchard Jr., Lykken, McGue, Segal, & Tellegen, 1990, p. 228). The idea is that genetic factors are expressed "by influencing the character, selection, and impact of experiences during development." Of course, environments here refer to the normal range of settings that support human development; unusual or extreme environments that reduce or deprive individuals of emotional and/or physical sustenance can leave enduring effects on intelligence (Turkheimer, Haley, Waldron, d'Onofrio, & Gottesman, 2003; Segal & Montoya, 2018).

The CTA-MZ's nearly identical mean IQ scores at Time 1 to Time 2 may not be so surprising. While children show increased intellectual capacities from ages 2 to 12 years (Volkova, 2014), they change in ways that are generally consistent with their peers. Of course, none of the reported correlations approached 1.0. An association between earlier age at adoption and higher IQ was observed. This association is sometimes attributed to the relatively good health of an infant who appears more favorable to prospective adoptive parents. However, parents adopting children from China do not choose a child, but are assigned a child prior to their arrival. It is plausible that earlier age at adoption allows for earlier access to the adoptive family's resources and reduced time in the orphanage, but this explanation requires further study with pairs of MZ cotwins adopted at different times. A final point is that, in contrast with age at adoption, age at separation did not correlate with the twins' IQ scores. Age at separation and age at adoption are different (albeit, related) measures. Age at separation may arise if one twin is kept by the family, the twins are placed in different temporary foster care situations and adopted separately, or other circumstances. In contrast, age at adoption (in the present study) is the child's age upon entry into the adoptive family, not the age of legal adoption. Lastly, contact during the Time 1 to Time 2 interval had no effect on young twins' IQ resemblance, weakening critics' assertion that such factors affect intellect in meaningful ways; see Farber (1981) and Segal (2012).

Table 4
Heritability estimates for Wechsler IQ data for monozygotic twins reared apart and environmentality.

Estimate for Virtual Twins								
Study	*N* (pairs)	Time 1[a]	Time 2[b]	Mean of *Rs*	*R* of Means[c]	R_{wt}	R_{ct}	R_{ct}/R_{wt}
CTA-MZ	15/14	0.51*	0.81***	0.65	0.69**	0.82***	0.57***	0.70
D-MZA	12/9	0.64**	0.74**	0.68	0.76**	0.88***	0.66***	0.75
VT	43	0.30*	0.11	0.21	0.18	0.50**	0.06	.12

CTA-MZ: Chinese MZ twins reared apart; D-MZA: Danish MZ twins reared apart; VT: Virtual twins.
R_{wt}: Retest; R_{ct}: Cross Time-Cross Twin IQ Correlation; R_{ct}/R_{wt}: Estimate of Set-Point Heritability.
* p <.05.
** p <.01.
*** p < .001.
[a] Intraclass correlation, CTA-MZ: $N = 15$ pairs; D-MZA: $N = 12$ pairs; VT: $N = 43$ pairs.
[b] Intraclass correlation, CTA-MZ: $N = 14$ pairs; D-MZA: $N = 9$ pairs; VT: $N = 43$ pairs.
[c] Correlation of mean IQ scores.

Changes in IQ similarity among the adult Danish twins were slight, but in the predicted directions. The significant increase in IQ score at Time 2 is worth noting. Only nine complete pairs participated in the second phase of the study. Inspection of IQ scores of the three pairs present at Time 1, but not at Time 2, did not suggest that they were atypical in any way; however, some pairs with higher Time 1 IQ scores remained, possibly explaining the mean IQ increase at Time 2. Unfortunately, only age at separation, but not age at adoption, was provided in that study, but showed no relation to IQ performance as in the CTA-MZ analysis.

The VTs' results were discussed previously and they are summarized briefly above. Again, developmental trends reflected reduced IQ resemblance and increased within-pair difference. These findings most likely reflect the effects of new genetic factors and/or nonshared environments, as well as the declining effects of the shared experiences of same-age unrelated siblings growing up together. These results depart from those of the twins in the present study, especially the CTA-MZ pairs.

4.2. Part II findings

Consistent with Lykken's (2007) paper, the estimates of IQ heritability for the CTA-MZ twins, based on two IQ measures, exceeded the estimates based on the *R* of Means. This reflects the greater stability of the repeated data than the single time scores, although additional longitudinal measures would be desirable. When IQ heritability was estimated in this way for the D-MZA twins, the *R* of Means and ratio (R_{ct}/R_{wt}) were nearly the same. This difference between the two samples most likely speaks to the greater continuity of the adult IQ scores from Time 1 to Time 2. For VTs, the reduced IQ resemblance over time is a likely result of the waning effects of shared environmental influences on IQ, consistent with the ordinary adoptive sibling studies cited above.

4.3. Limitations

Limitations to the present study are the small sample sizes that urge careful interpretation of the findings. Nevertheless, confirmation of the hypotheses specified at the outset is encouraging. It is fortunate that continued IQ analyses are planned using participants in the Minnesota Study of Twins Reared Apart who have been IQ tested on two occasions (Segal, 2012). Further analyses of the CTA-MZ twins are also planned and will, hopefully, identify additional participants for study.

4.4. Implications

The findings reported here should offer parents and educators insights into the academic performances and interests of twins and adoptees. MZ twins can be expected to achieve similar results on school tests, whereas unrelated siblings can be expected to achieve different outcomes. Knowing this will help parents and educators tailor their treatment, resource provision, and expectations of different children within families. In doing so, they may avoid the frustration that may come from encouraging and/or expecting outcomes and goals that may be outside the child's inclinations. Even parents and teachers of nontwin children can benefit from acknowledging the variety of influences that shape general intelligence.

CRediT authorship contribution statement

Nancy L. Segal: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing. **Elizabeth Pratt-Thompson:** Data curation, Writing – review & editing.

Declaration of competing interest

The authors have nothing to declare.

Data availability

The data that have been analyzed are confidential.

Acknowledgments

Dr. Tony Vernon, Mt. Royal University, critiqued an early version of this paper. The study was variously funded by the National Science Foundation (SBR-9712875), a grant from the National Institute of Health (NIMH-R01 MH63351) and intramural research grants from California State University, Fullerton to the senior author.

References

Bouchard, T. J., Jr., Lykken, D. T., McGue, M., Segal, N. L., & Tellegen, A. (1990). Sources of human psychological differences: The Minnesota study of twins reared apart. *Science, 250*(4978), 223–228.
Bouchard, T. J. (2013). The Wilson effect: The increase in heritability of IQ with age. *Twin Research and Human Genetics, 16*(5), 923–930.
Buckley, C. (29 October, 2015). "China ends one-child policy, allowing families two children." New York Times, https://www.nytimes.com/2015/10/30/world/asia/china-end-one-child-policy.html.
Deary, I. J., Whalley, L. J., Lemmon, H., Crawford, J. R., & Starr, J. M. (2000). The stability of individual differences in mental ability from childhood to old age: Follow-up of the 1932 Scottish mental survey. *Intelligence, 28*(1), 49–55.
Derner, G. F., Aborn, M., & Canter, A. H. (1950). The reliability of the Wechsler-Bellevue subtests and scales. *Journal of Consulting Psychology, 14*(3), 172–179.
Evans, K. (2008). *The lost daughters of China: Adopted girls, their journey to America, and the search for a missing past.* New York: TarcherPerigee.
Farber, S. L. (1981). *Identical twins reared apart.* NY: Basic Books.
Johnson, W., Bouchard, T. J., Jr., McGue, M., Segal, N. L., Tellegen, A., Keyes, M., & Gottesman, I. I. (2007). Genetic and environmental influences on the verbal-perceptual-image rotation (VPR) model of the structure of mental abilities in the Minnesota study of twins reared apart. *Intelligence, 35*(6), 542–562.
Loehlin, J. C., Horn, J. M., & Willerman, L. (1989). Modeling IQ change: Evidence from the Texas Adoption Project. *Child Development, 60*(4), 993–1004.
Lykken, D. T. (2007). A more accurate estimate of heritability. *Twin Research and Human Genetics, 10*(1), 168–173.
Lykken, D.T., Geisser, S., & Tellegen, A. (1981). Heritability estimates from twin studies: The efficiency of the MZA design. (unpublished manuscript).
McCartney, K., Harris, M. J., & Bernieri, F. (1990). Growing up and growing apart: A developmental meta-analysis of twin studies. *Psychological Bulletin, 107*(2), 226–237.
McClearn, G. E., Johansson, B., Berg, S., Pedersen, N. L., Ahern, F., Petrill, S. A., & Plomin, R. (1997). Substantial genetic influence on cognitive abilities in twins 80 or more years old. *Science, 276*(5318), 1560–1563.
McGuire, S. A., & Segal, N. L. (2013). Peer networks: Overlap in twin, sibling and friend dyads. *Child Development, 84*(2), 500–511.
Nichols, R. C., & Bilbro, W. C., Jr. (1966). The diagnosis of twin zygosity. *Acta Genetica et Statistica Medica, 16*(3), 265–275.
Sattler, J. M. (1989). *Assessment of children's intelligence* (2nd ed.). Philadelphia: W.B. Saunders.
Scarr, S., Weinberg, R. A., & Waldman, I. D. (1993). IQ correlations in transracial adoptive families. *Intelligence, 17*(4), 541–555.
Segal, N. L. (1997). Twin research perspective on human development. In N. L. Segal, G. E. Weisfeld, & C. C. Weisfeld (Eds.), *Uniting psychology and biology: Integrative perspectives on human development* (pp. 145–173). Washington, D.C.: APA Press.
Segal, N. L. (2006). Fullerton Virtual Twin Study. *Twin Research and Human Genetics, 9*(6), 963–964.
Segal, N. L. (2012). *Born together-reared apart: The landmark Minnesota Twin Study.* Cambridge, MA: Harvard University Press.
Segal, N. L., McGuire, S. A., Havlena, J., Gill, P., & Hershberger, S. L. (2007). Intellectual similarity of virtual twin pairs: Developmental trends. *Personality and Individual Differences, 42(7)*, 1209–1219.
Segal, N. L., & Montoya, Y. S. (2018). *Accidental brothers: The story of twins exchanged at birth and the power of nature and nurture.* New York: St: Martin's Press.
Segal, N. L., & Niculae, F. J. (2019). Fullerton Virtual Twin Project: Overview and 2019 update. *Twin Research and Human Genetics, 22*(6), 731–734.
Segal, N. L., Niculae, F. J., Becker, E. N., & Shih, E. Y. (2021). Reared-apart/reared-together Chinese twins and virtual twins: Evolving research program and general intelligence findings. *Journal of Experimental Child Psychology, 207*, 105106.
Segal, N. L., & Russell, J. (1991). IQ similarity in monozygotic and dizygotic twin children: Effects of the same versus separate examiners: A research note. *Journal of Child Psychology and Psychiatry, 32*(4), 703–708.
Segal, N. L., Tan, T. X., & Graham, J. L. (2015). Twins and virtual twins: Do genetic (as well as experiential) factors affect developmental risks? *Journal of Experimental Child Psychology, 136*, 55–69.

Turkheimer, E., Haley, A., Waldron, M., d`Onofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science, 14*(6), 623–662.

Volkova, E. V. (2014). Age dynamics of intelligence in adolescence and early adulthood. *Procedia-Social and Behavioral Sciences, 140*, 440–446.

Wilson, R. S. (1978). Synchronies in mental development: An epigenetic perspective: Early mental development of twins and siblings is analyzed within the framework of evolutionary theory. *Science, 202*(4371), 939–948.